



Munich Personal RePEc Archive

Underground Employment in the Regions of Italy: A Panel Analysis

Lisi, Gaetano
University of Cassino

13. November 2009

Online at <http://mpa.ub.uni-muenchen.de/18525/>
MPRA Paper No. 18525, posted 13. November 2009 / 12:12

Il lavoro irregolare in Italia: un'analisi panel regionale

Gaetano Lisi *

Novembre, 2009

Abstract

Il lavoro indaga empiricamente due importanti relazioni, oggetto di ampio interesse e di dibattito aperto in letteratura: il legame tra occupazione irregolare e disoccupazione e quello tra economia sommersa e contesto ambientale. I principali risultati dell'analisi sono due. Il primo: il rapporto di causalità che lega l'occupazione irregolare e la disoccupazione sembra essere bidirezionale, sebbene asimmetrico. Il secondo: le variabili "ambientali" e gli effetti non osservabili sono cruciali nella spiegazione delle differenze riscontrate a livello regionale nei tassi di irregolarità italiani.

JEL Classification: E26, O17, C23

Keywords: lavoro irregolare, modelli panel, economia sommersa

* Università di Cassino. Email: gaetano.lisi@unicas.it.

1. Introduzione

Il dualismo economico-sociale esistente in Italia e conosciuto in tutto il mondo (la c.d. “questione meridionale”), si manifesta nella sua forma più eclatante con riferimento all’economia sommersa. L’Italia, infatti, è uno dei Paesi più industrializzati con la quota più alta di economia sommersa; inoltre, nella medesima struttura istituzionale convivono due macro-aree caratterizzate da differenti livelli di occupazione irregolare (cfr. ISTAT, 2005, 2008).

Nelle regioni del Sud Italia, in cui più elevata è la quota di attività irregolari, si registrano anche i livelli più alti di disoccupazione (cfr. Boeri e Garibaldi, 2002).

La letteratura (sia empirica che teorica) è concorde nel ritenere che l’occupazione irregolare e la disoccupazione siano fenomeni strettamente connessi ed interdipendenti. Boeri e Garibaldi trovano conferma di ciò al punto che i due fenomeni sono da loro descritti come «[...] *due lati della stessa medaglia*» (2006, p. 20).

Controverso, tuttavia, è il legame di causalità esistente (cfr. Tanzi, 1999; Giles e Tedds, 2002). Ciò dipende dal fatto che la forza lavoro irregolare è notevolmente eterogenea. Infatti, una quota della forza lavoro irregolare è spesso erroneamente classificata tra i disoccupati ed è, quindi, annoverata tra la forza lavoro ufficiale; un’altra parte dei lavoratori irregolari, invece, è costituita da pensionati, minorenni, casalinghe, immigrati regolari e clandestini, che non fanno parte della forza lavoro inclusa nelle statistiche ufficiali. In aggiunta, ci sono individui che, allo stesso tempo, fanno parte sia della forza lavoro ufficiale che non ufficiale, il c.d. fenomeno del doppio lavoro (cfr. Tanzi, 1999).

Nei lavori empirici dedicati all’economia sommersa è prassi far uso del *model approach* o MIMIC (*Multiple Indicators Multiple Causes*) *method*, una metodologia di stima che tratta il sommerso come una variabile latente in funzione di una serie di indicatori e di determinanti (i.e. cause). Tuttavia, tale metodologia di stima produce solo un indice temporale per la variabile latente e il suo uso può essere sconsigliato dal fatto che l’economia sommersa non è propriamente una variabile non osservabile (Breusch, 2005).

In aggiunta, i lavori empirici che fanno uso anche di dati *panel* o *cross-section* focalizzano l'analisi sul confronto tra Paesi diversi (*between country analysis*), rendendo oltremodo difficile trarre conclusioni precise visto che i differenti contesti culturali ed istituzionali possono influenzare non poco il livello esistente di economia sommersa (cfr. Schneider e Torgler, 2007). Uno dei pochi lavori che ha fatto uso di un'analisi *panel within country* è quello di Schneider e Torgler (2007). Tale studio, però, si concentra sui cantoni svizzeri, dotati di ampia autonomia e, dunque, potenzialmente molto eterogenei dal punto di vista istituzionale.

L'interesse per la realtà italiana deriva proprio dal fatto che in un contesto istituzionale omogeneo convivono due macro aree differenti dal punto di vista economico-sociale. La variabile dipendente usata nella presente analisi è il tasso di irregolarità regionale che l'ISTAT stima in modo accurato e che permette di concentrare l'attenzione sul mercato del lavoro, visto che « *il lavoro è il principale fattore produttivo su cui si basa il funzionamento dell'economia sommersa* » (Lucifora, 2003, p. 79) e la decisione di operare nel sommerso è sostanzialmente il risultato di scelte fatte da lavoratori ed imprese che interagiscono nel mercato del lavoro (Bouev, 2005).

Sebbene la rilevanza dei fattori economico-istituzionali (quali la regolamentazione e il "peso" fiscale) nella spiegazione dell'economia sommersa sia una tesi ampiamente condivisa in letteratura (cfr. Schneider e Enste, 2000), è altrettanto condivisa l'idea che la corruzione, la *tax morality* (o *tax morale*)¹ e la bassa qualità delle istituzioni siano fattori altrettanto cruciali (Tanzi, 1988; Johnson et al., 2000; Sarte, 2000; Fugazza e Jacques, 2004; Schneider, 2007; Schneider e Torgler, 2007).

Inoltre, con riferimento al contesto italiano, non occorre dimenticare il ruolo imponente e negativo svolto dalla criminalità organizzata nel Mezzogiorno (cfr. Daniele e Marani, 2008; Marini e Turato, 2002).²

¹ L'atteggiamento degli individui nei confronti dello Stato, cioè la maggiore o minore predisposizione a rispettarne le regole (cfr. Schneider, 2007; Schneider e Torgler, 2007).

² Tale discorso può essere facilmente esteso anche al contesto europeo, visto che la corruzione e la criminalità sono particolarmente diffuse nei Paesi dell'Europa dell'Est, dove l'economia sommersa è più alta che nel resto d'Europa (cfr. Van Dijk, 2006; Johnson et al., 2000).

In particolare, nella presente analisi, oltre a specifiche variabili quali la corruzione, la criminalità, la dotazione infrastrutturale, l'istruzione, l'industrializzazione, si prenderanno in considerazione anche gli effetti specifici (tipici dei modelli panel), al fine di tener conto di tutte quelle variabili non osservabili (come il fattore storia, cultura e le esternalità) ugualmente importanti nella spiegazione del differente livello di economia sommersa esistente tra il *Nord* e il *Sud* d'Italia.

Il resto del lavoro è così organizzato: il secondo e il terzo paragrafo presentano, rispettivamente, il dataset usato e il modello econometrico scelto; nel quarto, i risultati dell'analisi empirica sono mostrati e discussi; infine, il quinto conclude.

2. Il dataset

Il campione di dati è costituito da un panel riguardante le 20 regioni d'Italia con 11 osservazioni temporali comprese tra il 1995 ed il 2005, per un totale di 220 osservazioni. Più precisamente, il panel utilizzato è composto da 11 variabili (cfr. tabella 1) rilevate nelle 20 regioni italiane per 11 periodi.

La variabile dipendente è il *tasso di irregolarità (LS)*, calcolato come rapporto percentuale tra le unità di lavoro irregolari di una regione e il complesso delle unità di lavoro occupate nella stessa area territoriale (cfr. tabella 2).³

La scelta delle variabili esplicative da includere nell'analisi è stata suggerita dal fatto, ormai consolidato, che la dimensione del sommerso, oltre a dipendere da fattori economico-istituzionali, dipende anche dal contesto ambientale in cui imprese e lavoratori si trovano ad operare.

Seguendo Daniele e Marani (2008), come *proxy* della dimensione e della struttura economica regionale si introducono due variabili: l'indice sintetico di dotazione infrastrutturale (*INFR*) e il tasso di industrializzazione (*IND*).

Al fine di cogliere l'incidenza della variabile "criminalità organizzata" (*IOCR*), è stato inserito nel modello un indice dato dalla somma dei reati tipici delle

³ Le unità di lavoro (ULA) rappresentano una misura di quanto il fattore lavoro contribuisca alla produzione del Paese in un determinato periodo. Sono calcolate attraverso la trasformazione ad unità a tempo pieno delle posizioni lavorative ricoperte da ciascuna persona occupata nel periodo di riferimento (per ulteriori dettagli, cfr. ISTAT, 2005, 2008).

organizzazioni di tipo mafioso denunciati (i.e. estorsioni e associazioni a delinquere) ogni 10.000 abitanti.

Per quanto riguarda la variabile “corruzione” (*ICO*), si è costruito un indice dato dalla somma delle sentenze di condanna per corruzione, peculato, abuso di ufficio e concussione ogni 10.000 abitanti.⁴

Tra le variabili di controllo si includono quelle usualmente annoverate tra le determinanti e/o indicatori del sommerso (cfr. Amendola e Dell’Anno, 2008; Dell’Anno, 2003; Dell’Anno e Schneider, 2003), i.e. il tasso di crescita del PIL pro-capite regionale (*GROWTHPIL*), il tasso di disoccupazione (*U*), il tasso di crescita delle entrate fiscali (*GROWHTAX*) e un indice relativo al peso della regolamentazione (*REG*).⁵ Le ultime due variabili incluse nel modello sono *individual-invariant* e hanno lo scopo di sintetizzare il contesto economico-istituzionale comune a tutte e 20 le regioni.

Vengono introdotte, infine, due variabili strettamente legate al sommerso: la composizione percentuale dell’entità dell’evasione IRAP (Imposta Regionale sulle Attività Produttive), come *proxy* della *tax morality* (*EF*), e il tasso di istruzione (*ISTR*). Per quanto riguarda la prima, l’intuizione sottostante è la forte correlazione (negativa) esistente tra *tax morality* ed evasione fiscale (cfr. Schneider, 2007; Schneider e Torgler, 2007); mentre, per quanto riguarda l’istruzione, il suo ruolo chiave nella riduzione del lavoro sommerso è confermato da recenti studi empirici (cfr. Cappariello e Zizza, 2009).

3. Il modello econometrico

Le analisi basate su dati panel sono un potente (se non il più potente) strumento di indagine statistico-econometrica. La combinazione dell’informazione *cross-section* (longitudinale) con quella *time-series* (temporale), infatti, permette di formulare e

⁴ Poiché i distretti di Corte d’Appello sono per lo più territorialmente riferibili alla ripartizione delle regioni amministrative (con alcune eccezioni superabili attraverso specifici accorpamenti) è stato possibile creare una mappatura regionale del fenomeno in considerazione.

⁵ Per la variabile *REG* la disponibilità di dati è quinquennale. Si è proceduto, quindi, ad un loro estensione: ad esempio, il dato del 1995 è stato usato anche per il 1996, 1997, 1998, 1999 fino alla disponibilità del nuovo dato nel 2000.

stimare modelli interpretativi migliori dei fenomeni economici oggetto di studio (Baltagi e Griffin, 1984).

I modelli panel possono avere diverse specificazioni. La maggior parte delle applicazioni economiche tendono ad essere caratterizzate da un grande numero di osservazioni individuali e da poche osservazioni temporali (Baltagi, 2008). In questi casi, le tecniche tendono a concentrarsi sulla variabilità individuale (Golinelli e Bontempi, 2007). Precisamente, nel caso in cui (come nello specifico dell'analisi proposta) $N > T$ e T non è sufficientemente grande per poter fare inferenza sui singoli individui, è prassi comune adottare la scelta di un'unica pendenza per tutti gli individui (Golinelli e Bontempi, 2007).

Nei modelli panel, il termine di errore $\varepsilon_{i,t}$ può essere scomposto in $\mu_i + \tau_t + v_{i,t}$, dove μ_i è l'effetto specifico individuale non osservabile, τ_t è l'effetto specifico temporale non osservabile, e $v_{i,t}$ denota il tradizionale disturbo stocastico. Se $\varepsilon_{i,t} = \mu_i + v_{i,t}$ si parla di modelli panel a "componente di errore ad una via" (*One-way Error Component Regression Model*), viceversa, se $\varepsilon_{i,t} = \mu_i + \tau_t + v_{i,t}$ si parla di modelli panel a "componente di errore a due vie" (*Two-way Error Component Regression Model*).⁶ In entrambi i casi, ciò che conta è l'assunzione circa la natura degli effetti specifici non osservabili. Se μ_i e τ_t sono ipotizzati deterministici, si parla di modello panel ad "effetti fissi" (*Fixed Effect Model*); se, invece, sono ipotizzati stocastici, si parla di modello panel ad "effetti casuali" (*Random Effect Model*).

Per Baltagi (2008), il modello ad effetti fissi è un'adeguata specificazione se l'analisi è focalizzata su un insieme specificato di N unità e l'inferenza è indirizzata all'andamento delle medesime unità; viceversa, il modello ad effetti casuali è la specificazione più adeguata se le N unità osservate sono casualmente selezionate dalla popolazione e si vuole inferire dal campione all'intera popolazione.

⁶ Nelle applicazioni panel caratterizzate da poche osservazioni temporali, è prassi utilizzare la più semplice specificazione *one-way* (Baltagi, 2008).

Inoltre, se gli effetti non osservabili rappresentano anche variabili omesse, è altamente probabile che queste caratteristiche *cross-country* siano correlate con gli altri regressori del modello, rendendo quindi indispensabile il ricorso ad un modello ad effetti fissi (Judson e Owen, 1999).⁷

Tuttavia, nello specifico della presente analisi, c'è un'altra importante motivazione che spinge all'uso di un modello panel ad effetti fissi: la "ricchezza" derivante dall'esistenza di una correlazione tra le variabili incluse nel modello e gli effetti fissi. Tale correlazione, infatti, potrebbe essere spiegata dall'esistenza di esternalità (sia positive che negative) prodotte da variabili quali l'istruzione, il livello di infrastrutture, la corruzione, la criminalità, ecc.

Il modello econometrico scelto risulta, dunque, essere il seguente:⁸

$$y_{i,t} = \sum_{j=1}^N \mu_i \cdot D_{i,j} + x'_{i,t} \beta + v_{i,t}$$

dove $D_{i,j} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$ sono le *dummies* incluse per catturare gli effetti fissi *cross-country*.

4. I risultati dell'analisi empirica

Tutte le variabili incluse nel modello sono espresse in logaritmi, al fine di ottenere dai coefficienti stimati delle misure di elasticità immediatamente confrontabili.⁹

La tabella 4 mostra i risultati della stima del modello ad effetti fissi in cui si è fatto uso dello stimatore *Least Squares Dummies Variables (LSDV)* corretto per la presenza

⁷ Poiché il modello ad effetti casuali tratta gli effetti non osservabili in modo stocastico, la differenza principale rispetto al modello ad effetti fissi è l'assunzione che tali effetti non siano correlati con i regressori del modello (per tutte le i e t).

⁸ Le scelte di specificazione effettuate in base ai suggerimenti presenti in letteratura vanno confermate attraverso opportuni test statistici. I risultati dei principali test eseguiti sono proposti in tabella 3.

⁹ In particolare $GROWTHPIL = \ln(PIL_t) - \ln(PIL_{t-1})$. Identico discorso per $GROWHTAX$. Per le altre variabili l'uso del logaritmo naturale è denotato da una lettera "l" posta all'inizio della notazione scelta (ad esempio, "lls" è il logaritmo naturale della variabile "ls").

di eteroschedasticità.¹⁰ Tale stimatore, a differenza di quello *Within*, consente di stimare anche le variabili *time-invariant* (nello specifico *INFR* e *EF*)¹¹ e di evidenziare tutti gli effetti fissi individuali, di cui occorre creare le *dummies*.¹² La stima *LSDV* è molto onerosa in termini di perdita di gradi di libertà (occorre, infatti, stimare $N - 1$ parametri aggiuntivi), ragion per cui le *dummies* regionali vengono accorpate per macro-aree geografiche, i.e. Nord-Est, Nord-Ovest, Centro e Mezzogiorno (secondo la suddivisione proposta dall'ISTAT).¹³

Al fine di testare la significatività degli effetti individuali, ovvero la correttezza di preferire un modello ad effetti fissi piuttosto che un modello *pooled* (in cui non solo le pendenze ma anche le intercette sono uguali per tutti gli individui), è possibile eseguire un *test F*: il *test* rigetta l'ipotesi nulla che tutti gli effetti fissi individuali siano uguali a zero (cfr. tabella 3), segnalando l'opportunità di usare un modello ad effetti fissi, i.e. evidenziando il ruolo significativo degli effetti individuali.¹⁴

In aggiunta, gli effetti fissi individuali oltre ad essere in blocco significativamente diversi da zero, sono significativi anche per macro-aree geografiche (cfr. i *t-test* associati a ciascun effetto), il che dovrebbe implicare anche l'assenza di rilevanti problemi di multicollinearità.

Il test di *Hausman* conferma, dal punto di vista econometrico, la bontà della scelta di ricorrere ad uno stimatore ad effetti fissi piuttosto che uno ad effetti casuali.

¹⁰ Assumendo che i disturbi della regressione siano omoschedastici, con la stessa varianza nel tempo e tra individui, quando in realtà l'eteroschedasticità è presente (principalmente, nel caso di panel, tra le unità *cross-section*), determinerà degli stimatori ancora consistenti ma inefficienti. In questo caso, gli standard errors saranno distorti e occorrerà correggerli per la presenza di eteroschedasticità (Baltagi, 2008, pag. 87). Inoltre, come nota Wooldridge (2002, pag. 274), mentre l'eteroschedasticità è sempre un potenziale problema, la correlazione seriale può esserlo in determinate applicazioni, soprattutto quando T è molto grande.

¹¹ A causa dell'indisponibilità di dati, per tali variabili si è potuto far uso solo di un indice *cross-section*.

¹² È semplice verificare che lo stimatore *LSDV* produce stime di β identiche a quelle prodotte dallo stimatore *Within* (cfr. Wooldridge, 2002). L'unica differenza, in sostanza, è nell'uso delle *dummies*: la costante riportata nella stima *Within*, infatti, altro non è che la media degli effetti individuali.

¹³ Tuttavia, per $N \rightarrow \infty$ si hanno stime inconsistenti dei soli effetti individuali (cfr. Baltagi, 2008; Wooldridge, 2002).

¹⁴ Dal momento che l'operazione di "pooling" consiste nell'applicare restrizioni lineari sui coefficienti, il *test F* verifica la validità del modello ristretto (i.e. il modello *pooled*) rispetto ad un modello non-ristretto dove non sono imposti vincoli sui coefficienti (i.e. il modello ad effetti fissi).

Per quanto riguarda le stime, l'unica variabile statisticamente non significativa è *GROWHTHAX*. Ciò non è sorprendente, visto che la relazione tra la percentuale di economia sommersa e la quota di gettito fiscale è meno semplice di quanto appare (cfr. Johnson et al., 1998, 1999; Friedman et al., 2000; Bovi e Dell'Anno, 2007). Inoltre, la variabile *IOCR* ha un impatto su *LS* superiore a quello della variabile *REG*. Questi due risultati giustificherebbe l'enfasi posta nella presente analisi sulle variabili "ambientali".

Non senza sorprese, il segno della relazione tra il tasso di crescita del PIL (*GROWTHPIL*) e l'economia sommersa è positivo e significativo. Riguardo a tale relazione, la letteratura empirica mostra una notevole eterogeneità nei risultati anche facendo uso della stessa metodologia di stima dell'economia sommersa (cfr. Eilat e Zinnes, 2000; Dell'Anno, 2003). Tale ambiguità è solo apparentemente sorprendente, tenuto conto che le attività irregolari pur essendo anticicliche contribuiscono al reddito e alla produzione nazionale (Busato e Chiarini, 2004).

Attesi e significativi sono i segni delle variabili *ICO* (+), *INFR* e *IND* (-), mentre dalla positività e significatività della variabile *EF* si deduce che un aumento della *tax morality* riduce *LS* (come in Schneider e Torgler, 2007).

Va segnalato, infine, il ruolo chiave giocato dall'istruzione (*ISTR*) nel ridurre l'occupazione irregolare (come in Cappariello e Zizza, 2009).

4.1 Come indagare il legame di causalità tra *LS* e *U* ?

Tenuto conto dello stretto e ambiguo legame di causalità esistente, le variabili *LS* e *U* sono considerate endogene. Nulla, dunque, può essere dedotto dalla precedente analisi circa tale legame.

Al fine di superare il problema dell'endogeneità dei regressori, la letteratura empirica suggerisce il ricorso a modelli con variabili strumentali (cfr. Hausman e Taylor, 1981; Amemiya e MaCurdy, 1986; Breusch, Mizon and Schmidt, 1989). Tali procedure, però, sono soggette a problematiche non banali, quali l'individuazione di strumenti idonei e la loro efficacia. Infatti, la difficoltà (se non l'impossibilità) di individuare strumenti credibili (dal punto di vista della teoria economica) e

adeguati (dal punto di vista econometrico) per LS e U , impedisce di fatto il ricorso a tali modelli.

Una possibile alternativa è quella di poter strumentare le variabili con i loro ritardi (cfr. Anderson e Hsiao, 1981; Arellano, 1989; Arellano e Bond, 1991), senza necessariamente passare ad un modello dinamico. Se gli errori non sono autocorrelati,¹⁵ infatti, qualsiasi variabile $x_{i,t-s}$, con $s \geq 1$, risulterà essere uno strumento valido di $x_{i,t}$.¹⁶

Data la ridotta dimensione temporale del *dataset*, si limiterà l'uso di strumenti. Tuttavia, in presenza della fortissima persistenza dei due fenomeni, la differenza tra l'uso di 3 o 20 strumenti (ritardi) è pressoché irrilevante.

Tenuto conto dell'endogeneità di LS e U , i.e. della loro reciproca dipendenza, si fa ricorso per la stima ad un modello strutturale, i.e. alla c.d. procedura in tre fasi (*3SLS – three Stages Least Square*).¹⁷

La stima congiunta di LS e U , riportata nella tabella 5, offre un risultato molto interessante: il tasso di disoccupazione influenza (positivamente) il sommerso e viceversa. I coefficienti, infatti, sono entrambi significativi e con segno positivo.¹⁸ La relazione di causalità che lega il tasso di disoccupazione e il tasso di irregolarità sembra, dunque, essere bidirezionale. Tale relazione, inoltre, risulta essere asimmetrica, visto che l'impatto di U su LS è più "forte" di quello di LS su U .

¹⁵ Nello specifico, il test di Wooldridge non rifiuta l'ipotesi nulla di assenza di correlazione seriale $AR(1)$.

¹⁶ Affinché uno strumento sia valido devono valere le seguenti ipotesi: *i*) $E(x_{i,t}, x_{i,t-s}) \neq 0$, con $s \geq 1$, i.e. lo strumento deve essere correlato con la variabile da strumentare; *ii*) $E(x_{i,t-s}, \varepsilon_{i,t}) = 0$, i.e. lo strumento non deve essere correlato con l'errore. Ovviamente maggiore è il valore di T , maggiore è il numero di strumenti utilizzabili. Inoltre, condizione (necessaria) di identificazione dei parametri è che il numero di strumenti esogeni sia uguale o superiore al numero di variabili endogene da strumentare.

¹⁷ Non esistono, al momento, routine in STATA in grado di eseguire la procedura 3SLS con dati panel. L'alternativa è l'uso della funzione STATA: *xtdata, fe*. Per maggiori dettagli circa l'uso di tale routine si consulti l'help del software.

¹⁸ Una semplice spiegazione a questo risultato può essere la seguente: un aumento del tasso di disoccupazione implica un aumento dei posti di lavoro irregolari poiché, intuitivamente, ciò comporta un aumento della manodopera a disposizione delle imprese che assumono in nero; viceversa, un aumento dei posti di lavoro irregolari non riduce il tasso di disoccupazione poiché difficilmente coloro che trovano un lavoro irregolare si dichiareranno occupati irregolari. Inoltre, un aumento dei posti di lavoro irregolari può implicare una riduzione di quelli regolari e dunque dei lavoratori "ufficiali", con il conseguente aumento dei disoccupati ufficiali.

In aggiunta, alcuni ulteriori considerazioni possono essere fatte: *a)* Le variabili “corruzione” (*ICO*) e “criminalità organizzata” (*IOCR*) sono significative solo con riferimento alla dipendente *LS*. Questo risultato, da un lato conferma la mancanza di conclusioni univoche nella letteratura empirica che si occupa del legame tra disoccupazione e criminalità (cfr. Marselli e Vannini, 2000),¹⁹ dall’altro evidenzia il forte legame di tali fattori con l’estensione del settore sommerso (cfr. Van Dijk, 2006); *b)* Il tasso di crescita del *PIL* è significativo solo con riferimento alla dipendente *U* (ovviamente con segno negativo). Il risultato di un segno positivo tra il tasso di crescita del *PIL* e *LS* non è, dunque, robusto e ciò conferma l’ambiguità presente in letteratura circa il segno di tale relazione; *c)* Le variabili *GROWTHTAX* e *REG* sono statisticamente significative (con segno positivo) solo con riferimento alla variabile *U*, mentre l’istruzione (*ISTR*) è statisticamente significativa solo nella riduzione dell’occupazione irregolare; *d)* Infine, dal punto di vista della *policy*, tenuto conto della relazione di causalità che sembrerebbe esistere, un provvedimento che produce un effetto diretto su di una variabile potrebbe implicare anche un effetto indiretto sull’altra.

5. Conclusioni

I principali risultati del lavoro sono i seguenti:

- i.* La relazione di causalità che lega il tasso di disoccupazione e il tasso di irregolarità sembra essere bidirezionale. In aggiunta, tale relazione risulta essere asimmetrica, visto che l’impatto di *U* su *LS* è più “forte” di quello di *LS* su *U*.
- ii.* In aree caratterizzate da elevati livelli di corruzione e criminalità, il ruolo svolto da fattori quali la tassazione e la regolamentazione non sembra essere cruciale nella spiegazione dei differenti tassi di irregolarità riscontrati a

¹⁹ Occorre sottolineare, però, che nello specifico si fa riferimento alla sola criminalità organizzata e non anche alla microcriminalità. Ragion per cui tale ambiguità risulta essere ulteriormente accentuata nella presente analisi.

livello regionale; viceversa, importante è il ruolo svolto dal contesto ambientale (*tax morality* compresa).

- iii.* I fattori non osservabili specifici di ogni regione sono statisticamente significativi nella spiegazione dei differenti tassi di irregolarità riscontrati a livello regionale.
- iv.* Importante è anche il ruolo giocato dall'istruzione nel ridurre l'occupazione irregolare. In generale, una maggiore istruzione dovrebbe implicare maggiori e migliori opportunità di lavoro.

APPENDICE

Tabella 1. Descrizione delle variabili e delle fonti

<i>Variabile</i>	<i>Descrizione</i>	<i>Fonte</i>
LS	Tasso di irregolarità	ISTAT
GROWTHPIL	Tasso di crescita del PIL regionale pro-capite	Elaborazione su dati ISTAT
INFR	Indice di dotazione infrastrutturale regionale (indicatore Italia = 100)	Istituto Tagliacarne
IND	Occupati totali regionali industria in senso stretto (media annua in migliaia) in rapporto agli occupati totali regionali (media annua in migliaia)	Elaborazione su dati ISTAT
ISTR	Composizione percentuale di laureati e diplomati sulla popolazione residente regionale di 15 anni e oltre	ISTAT
U	Tasso di disoccupazione	ISTAT
EF	Composizione percentuale dell'entità dell'evasione IRAP (come <i>proxy</i> della <i>tax morality</i>)	Agenzia delle Entrate
ICO	Somma delle sentenze di condanna per corruzione, peculato, abuso di ufficio e concussione ogni 10.000 abitanti (indice successivamente moltiplicato per 100)	Elaborazione su dati Casellario Giudiziale Centrale
IOCR	Somma dei reati tipici delle organizzazioni criminali denunciati (estorsioni e associazioni a delinquere, comprese quelle di tipo mafioso) ogni 10.000 abitanti (indice successivamente moltiplicato per 100)	Elaborazione su dati ISTAT ("sistema informativo territoriale sulla giustizia")
GROWTHTAX	Tasso di crescita delle entrate fiscali totali riscosse dallo Stato in percentuale del PIL	Elaborazione su dati OECD
REG	Indice di regolamentazione	OECD

Tabella 2. Tassi di irregolarità (*) - Totale economia (fonte: ISTAT, 2005, 2008)

Regioni	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Piemonte	10.3	10.8	10.6	10.2	10.5	11.1	10.8	9.5	8.3	8.8	9.7
Valle d'Aosta	16.0	15.3	15.7	17.5	16.3	15.5	10.0	9.9	9.8	10.6	10.8
Lombardia	11.4	11.2	11.1	10.9	10.5	10.2	9.4	8.1	7.0	7.6	7.8
Trentino Alto Adige	12.8	13.8	14.1	14.9	12.6	12.8	9.1	8.6	8.4	8.4	9.0
Veneto	11.2	11.0	10.9	11.4	11.1	11.5	9.9	8.8	7.9	8.3	8.7
Friuli Venezia Giulia	11.5	11.5	11.1	11.4	13.0	12.6	11.4	10.7	9.9	9.8	10.2
Liguria	12.8	13.5	13.3	13.8	13.7	13.2	14.0	12.2	10.9	11.7	12.5
Emilia Romagna	10.7	10.5	10.6	10.8	10.8	10.5	9.4	8.5	7.4	7.5	8.0
Toscana	11.9	12.3	12.9	12.9	12.9	13.0	10.6	9.5	8.4	8.4	9.0
Umbria	14.7	14.1	15.2	14.4	15.1	17.1	14.8	13.0	11.0	12.0	12.3
Marche	11.7	11.7	12.0	12.0	12.9	13.9	11.8	10.5	9.8	9.8	9.5
Lazio	16.5	16.5	17.0	17.2	16.9	17.2	15.1	13.1	11.0	12.1	11.9
Abruzzo	12.1	12.8	12.9	13.4	13.2	13.9	13.5	13.6	12.0	12.0	12.5
Molise	14.2	15.6	15.9	16.5	16.2	18.0	18.2	18.5	18.1	17.3	18.6
Campania	23.8	23.8	25.0	26.2	25.6	24.7	23.0	22.2	21.2	21.0	20.0
Puglia	19.4	19.5	19.4	19.4	19.5	20.4	18.8	18.2	16.9	15.5	16.4
Basilicata	17.1	17.5	18.1	19.9	19.8	22.1	19.0	19.3	19.8	18.7	20.1
Calabria	28.1	27.3	27.5	28.3	28.0	29.1	26.0	26.0	24.7	26.2	26.9
Sicilia	20.3	21.1	21.9	23.4	23.5	23.3	23.0	21.9	21.4	19.7	21.4
Sardegna	16.4	17.5	18.7	19.7	19.5	18.4	18.4	17.2	18.2	19.6	19.4
Totale Italia	14.5	14.5	14.8	15.1	15.0	15.0	13.8	12.7	11.6	11.7	12.1
<i>Nord-ovest</i>	11.3	11.3	11.2	11.0	10.8	10.8	10.2	8.9	7.7	8.3	8.8
<i>Nord-est</i>	11.2	11.1	11.1	11.5	11.3	11.3	9.8	8.9	8.0	8.2	8.6
<i>Centro</i>	14.2	14.2	14.8	14.9	14.9	15.4	13.1	11.5	10.0	10.5	10.7
<i>Mezzogiorno</i>	20.7	20.9	21.6	22.5	22.3	22.4	21.1	20.4	19.7	19.2	19.6

(*) incidenza % delle unità di lavoro non regolari sul totale delle unità di lavoro.

Tabella 3. Risultati dei principali test

Test eteroschedasticità

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: fitted values of lls

chi2(1) = 3.95

Prob > chi2 = 0.0470

Test autocorrelazione

lls

Wooldridge test for autocorrelation in panel data

H0: no first order autocorrelation

F(1, 19) = 2.401

Prob > F = 0.1421

lu

Wooldridge test for autocorrelation in panel data

H0: no first order autocorrelation

F(1, 19) = 4.448

Prob > F = 0.1841

Test di Hausman

b = consistent under Ho and Ha; obtained from xtreg

B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)

= 23.83

Prob>chi2 = 0.0024

Test F

(1) NordEst = 0.0

(2) NordOvest = 0.0

(3) Centro = 0.0

(4) Mezzogiorno = 0.0

F(4, 186) = 12.51

Prob > F = 0.0000

Tabella 4. Stima modello panel ad effetti fissi – stimatore LSDV

Regression with robust standard errors					Number of obs = 200		
					F(14, 186) = 976.59		
					Prob > F = 0.0000		
					R-squared = 0.8579		
					Root MSE = .32547		

	lls	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	

growthpil		.0292772	.0079662	3.68	0.000	.0135614	.044993
lu		.090462	.0364993	2.48	0.014	.0184636	.1624603
lef		.1205416	.0341058	3.53	0.001	.0532576	.1878256
linfr		-.0823364	.0388484	-2.12	0.035	-.1589765	-.0056963
lind		-.3045166	.0409019	-7.45	0.000	-.3852078	-.2238254
listr		-.2879223	.0342182	-8.41	0.000	-.355421	-.2204237
lico		.0301175	.0131165	2.30	0.023	.0042413	.0559937
liocr		.2334032	.050487	4.62	0.000	.1338024	.333004
lreg		.1780421	.0413203	4.31	0.000	.0965255	.2595587
growthtax		.0040793	.0046735	0.87	0.384	-.0051396	.0132983
NordEst		1.606557	.6791933	2.37	0.019	.2666447	2.94647
NordOvest		1.55469	.6771447	2.30	0.023	.2188185	2.890561
Centro		1.676228	.684617	2.45	0.015	.3256153	3.02684
Mezzogiorno		1.870692	.6498458	2.88	0.004	.5886757	3.152707

Tabella 5. Stima modello strutturale (3SLS) – stimatore Within

Three-stage least-squares regression						
Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
lls	160	9	.1047685	0.4251	1075.71	0.0000
lu	160	9	.107568	0.3150	534.65	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
lls						
lu	.162597	.0843126	1.93	0.054	-.0026526	.3278466
linfr	(dropped)					
lind	-.1431504	.1865485	-0.77	0.443	-.5087787	.2224779
listr	-.2298174	.06716	-3.42	0.001	-.3614486	-.0981862
lef	(dropped)					
lico	.0318668	.0109627	2.91	0.004	.0103804	.0533533
liocr	.2783391	.1137371	2.45	0.014	.0554185	.5012598
lreg	.047762	.057511	0.83	0.406	-.0649576	.1604815
growthpil	.0029588	.0067395	0.44	0.661	-.0102504	.016168
growthtax	.0056672	.0129191	0.44	0.661	-.0196537	.0309881
lu						
lls	.0460808	.0213409	2.16	0.031	.0042457	.0879158
linfr	(dropped)					
lind	-.1423829	.1674074	-0.85	0.395	-.4704953	.1857296
listr	-.1906086	.2987655	-0.64	0.523	-.7761783	.394961
lef	(dropped)					
lico	.04797	.1383794	0.35	0.729	-.2232487	.3191887
liocr	.1576188	.2111984	0.75	0.455	-.2563225	.5715601
lreg	.4523961	.1190232	3.80	0.000	.219115	.6856773
growthpil	-.0178086	.0101823	-1.75	0.080	-.0377655	.0021484
growthtax	.0249763	.0057887	4.31	0.000	.0136306	.036322

Endogenous variables: lls lu

Exogenous variables: growthpil linfr lind lef lico liocr growthtax lreg listr lls_1 lls_2 lls_3 lu_1 lu_2 lu_3

Riferimenti Bibliografici

- Amemiya, T. e MaCurdy, T. E. (1986). Instrumental Variable Estimation of an Error Components Models. *Econometrica*, 54, 869-881;
- Amendola, A., e Dell'Anno, R. (2008). Istituzioni, Disuguaglianza ed Economia Sommersa: quale relazione ? *Quaderno n. 24/2008*, Dipartimento di Scienze Economiche, Matematiche e Statistiche; Università degli Studi di Foggia;
- Anderson, T. W., e Hsiao, C. (1981). Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association*, 76, 598-606;
- Arellano, M. (1989). A Note on the Anderson-Hsiao Estimator for Panel Data. *Economics Letters*, vol. 31(4), 337-341;
- Arellano, M., e Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58, 277-297;
- Baltagi, B. H., e Griffin, J. M. (1984). Short and Long Run Effects in Pooled Models. *International Economic Review*, vol. 25(3), 631-45;
- Baltagi, B. H. (2008). *Econometric Analysis of Panel Data* (fourth edition). John Wiley and Sons, New York;
- Boeri, T., e Garibaldi, P. (2002). Shadow Activity and Unemployment in a Depressed Labour Market, *CEPR Discussion Paper*, n. 3433;
- Boeri, T., e Garibaldi, P. (2006). Shadow Sorting. *Fondazione Collegio Carlo Alberto Working Paper Series*, n. 10;
- Bovi, M., e Dell'Anno, R. (2007). The Changing Nature of the OECD Shadow Economy. *ISAE Working Paper*, n. 81;
- Breusch, T., Mizon, G., e Schmidt, P. (1989). Efficient Estimation Using Panel Data. *Econometrica*, 57, 695-700;
- Breusch, T. (2005). Estimating the underground economy using MIMIC models. *Working Paper University of Canberra, Australia*;
- Busato, F., e Chiarini, B. (2004). Market and Underground Activities in a Two Sector Dynamic Equilibrium Model. *Economic Theory*, 23(4): 831-861;
- Cappariello, R., and Zizza, R. (2009). *Dropping the Books and Working Off the Books*. Temi di discussione, *Working Papers*, Ufficio Studi Banca d'Italia, n. 702;
- Daniele, V., e Marani, U. (2008). Criminalità e investimenti esteri. Un'analisi per le province italiane. *Working Paper*, Università Magna Graecia di Catanzaro;
- Dell'Anno, R. (2003). Estimating the Shadow Economy in Italy: a Structural Equation Approach, *Working Paper No. 2003-07*, University of Aarhus;
- Dell'Anno, R., e Schneider, F. (2003). The Shadow Economy of Italy and Other OECD Countries: What do We Know ? *Journal of Public Finance and Public Choice*, XXI (2-3), 97-120;
- Eilat Y., e Zinnes, C. (2000). The Evolution of the Shadow Economy in Transition Countries: Consequences for Economic Growth and Donor Assistance. *CAER II Discussion Paper No. 83. Harvard Institute for International Development*;
- Friedman, E., Johnson, S., Kaufmann, D., e Zoido-Lobaton, P. (2000). Dodging the Grabbing Hand: the Determinants of Unofficial Activity in 69 Countries. *Journal of Public Economics*, 76, 459-493;
- Fugazza, M., e Jacques, F. (2004). Labour Market Institutions, Taxation and the Underground Economy. *Journal of Public Economics*, 88;
- Giles D. E. A., e Tedds, L. M. (2002). Taxes and the Canadian Underground Economy. *Canadian Tax paper n.106. Canadian Tax Foundation*;
- Golinelli, R., e Bontempi, M. E. (2007). *Panel Data Econometrics: Theory and Applications in STATA*, mimeo;

- Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica*, 46, 1251-1271;
- Hausman J., Taylor W., (1981). Panel Data and Unobservable Individual Effect. *Econometrica*, 49, 1377-1398;
- ISTAT (2005). *La misura dell'economia sommersa secondo le statistiche ufficiali*;
- ISTAT (2008). *La misura dell'occupazione non regolare nelle stime di contabilità nazionale*;
- Johnson, S., Kaufmann, D., e Zoido-Lobaton, P. (1998). Regulatory Discretion and the Unofficial Economy. *American Economic Review*, 88 (2);
- Johnson, S., Kaufmann, D., e Zoido-Lobaton, P. (1999). Corruption, Public Finances, and the Unofficial Economy. *Policy Research Working Paper Series*, 2169;
- Johnson, S., Kaufmann, D., McMillan, J., e Woodruff, C. (2000). Why do Firms Hide ? Bribes and Unofficial Activity after Communism. *Journal of Public Economics*, 76: 495-520;
- Judson, R. A., e Owen, A. L. (1999). Estimating Dynamic Panel Data Models: a Guide for Macroeconomists. *Economics Letters*, 65(1), 9-15 (October);
- Lucifora, C. (2003). *Economia sommersa e lavoro nero*. Il Mulino, Bologna;
- Marini D., e Turato, F. (2002). Nord-Est e Mezzogiorno: nuove relazioni, vecchi stereotipi. *Rapporti Formez-Fondazione Nord-Est*, aprile;
- Marselli, R., e Vannini, M. (2000). Quanto incide la disoccupazione sui tassi di criminalità ? *Rivista di Politica Economica*, 273-299.
- Sarte, P.D. G. (2000). Informality and Rent-seeking Bureaucracies in a Model of Long run Growth. *Journal of Monetary Economics*, Vol. 46, pp.173-197;
- Schneider, F., e Enste, D. H. (2000). Shadow Economies: Size, Causes and Consequences. *Journal of Economic Literature*, 38 (1), 77-114;
- Schneider, F. (2007). Shadow Economies and Corruption All Over the World: New Estimates for 145 Countries. *Economics e-Journal*, n. 2007-9 (July);
- Schneider, F., e Torgler, B. (2007). Shadow Economy, Tax Morale, Governance and Institutional Quality: A Panel Analysis. *Working Paper University of Linz*, n. 0701;
- Tanzi, V. (1988). Corruption around the World: Causes, Consequences, Scope and Cures. *IMF Working Paper*, Washington, D.C., n. 63;
- Tanzi, V. (1999). Uses and Abuses of Estimates of the Underground Economy. *The Economic Journal*, 109 (June), pp. 338-347;
- Van Dijk, J. (2006). *Organized Crime and Collective Victimization*. Paper for International Conference on Corruption and Organized Crime: Bridging Criminal and Economic Policies, June 23-24, 2006, Center for the Study of Democracy, Sofia, Bulgaria;
- Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.